

RESEARCH LETTER

10.1002/2016GL071794

Key Points:

- Methane and VOCs are used for source apportionment of methane emissions
- SFBA methane emissions are 1.3–2.3 times a recent inventory
- Biological source emissions are higher than fossil-fuel emissions

Supporting Information:

- Supporting Information S1

Correspondence to:

S. Jeong,
sjeong@lbl.gov

Citation:

Jeong, S., et al. (2017), Estimating methane emissions from biological and fossil-fuel sources in the San Francisco Bay Area, *Geophys. Res. Lett.*, 44, 486–495, doi:10.1002/2016GL071794.

Received 31 OCT 2016

Accepted 20 DEC 2016

Accepted article online 30 DEC 2016

Published online 9 JAN 2017

Estimating methane emissions from biological and fossil-fuel sources in the San Francisco Bay Area

Seongeun Jeong¹ , Xinguang Cui¹ , Donald R. Blake², Ben Miller³ , Stephen A. Montzka³ , Arlyn Andrews³ , Abhinav Guha⁴, Philip Martien⁴ , Ray P. Bambha⁵, Brian LaFranchi⁵ , Hope A. Michelsen⁵, Craig B. Clements⁶ , Pierre Glaize⁶ , and Marc L. Fischer¹ 

¹Energy Analysis and Environmental Impacts Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA,

²Department of Chemistry, University of California, Irvine, California, USA, ³Earth System Research Laboratory, NOAA,

Boulder, Colorado, USA, ⁴Bay Area Air Quality Management District, San Francisco, California, USA, ⁵Sandia National

Laboratories, Livermore, California, USA, ⁶Department of Meteorology and Climate Science, San Jose State University, San Jose, California, USA

Abstract We present the first sector-specific analysis of methane (CH₄) emissions from the San Francisco Bay Area (SFBA) using CH₄ and volatile organic compound (VOC) measurements from six sites during September – December 2015. We apply a hierarchical Bayesian inversion to separate the biological from fossil-fuel (natural gas and petroleum) sources using the measurements of CH₄ and selected VOCs, a source-specific 1 km CH₄ emission model, and an atmospheric transport model. We estimate that SFBA CH₄ emissions are 166–289 Gg CH₄/yr (at 95% confidence), 1.3–2.3 times higher than a recent inventory with much of the underestimation from landfill. Including the VOCs, 82 ± 27% of total posterior median CH₄ emissions are biological and 17 ± 3% fossil fuel, where landfill and natural gas dominate the biological and fossil-fuel CH₄ of prior emissions, respectively.

1. Introduction

With increased interest in mitigating rapid climate change impacts [e.g., *Hu et al.*, 2013], California's Short Lived Climate Pollutant Act (SB-605) includes a focus on methane (CH₄) emission mitigation [*California Legislative Information*, 2014]. Recent top-down studies in California [e.g., *Jeong et al.*, 2013; *Johnson et al.*, 2014; *Wecht et al.*, 2014; *Turner et al.*, 2015] suggest that California's CH₄ emissions are underestimated in the official greenhouse gas (GHG) inventory [*California Air Resources Board (CARB)*, 2014, 2015]. While both California Air Resources Board (CARB)'s inventory and top-down studies using atmospheric measurements [*Jeong et al.*, 2013, 2016; *Wecht et al.*, 2014; *Turner et al.*, 2015] suggest that California's Central Valley is the major source of the state's CH₄ emissions, increasing evidence suggests that California's large metropolitan regions constitute an important component of the state's anthropogenic CH₄ budget [*Hsu et al.*, 2010; *Wennberg et al.*, 2012; *Peischl et al.*, 2013; *Cui et al.*, 2015; *Fairley and Fischer*, 2015; *Wong et al.*, 2015; *Hopkins et al.*, 2016; *Wunch et al.*, 2016; *Jeong et al.*, 2016; *Johnson et al.*, 2016]. Furthermore, multiple recent studies in South Coast Air Basin (SoCAB; Los Angeles and surrounding areas) suggest that identifying CH₄ emission sources is important at the subregional or city scale for mitigation planning, but source partitioning is much more uncertain than estimating total CH₄ budget at this scale [*Wennberg et al.*, 2012; *Jeong et al.*, 2013, 2016; *Peischl et al.*, 2013; *Wong et al.*, 2015; *Cui et al.*, 2015; *Hopkins et al.*, 2016; *Conley et al.*, 2016].

Few studies have been conducted to estimate CH₄ emissions in the San Francisco Bay Area (SFBA), despite the fact that the total CH₄ emissions of this metropolitan region are estimated to be approximately 30% of CH₄ emissions from SoCAB, based on the CARB and Bay Area Air Quality Management District (BAAQMD) inventories [*CARB*, 2014; *Bay Area Air Quality Management District (BAAQMD)*, 2015]. Recent studies report that the central (i.e., mean) estimates for CH₄ emissions in SFBA range from 240 to 430 Gg/yr [*Fairley and Fischer*, 2015; *Jeong et al.*, 2016; *Wecht et al.*, 2014], which are larger than the BAAQMD inventory by factors of ~2–4. Given the discrepancy between the bottom-up inventory and the top-down analysis, it is important to further verify the estimates by *Fairley and Fischer* [2015], *Jeong et al.* [2016], and *Wecht et al.* [2014] and identify potential emissions sources for efficient mitigation planning, which has not been addressed by the previous studies.

We conducted a short-term measurement campaign to quantify CH₄ emissions from SFBA by measuring CH₄ and volatile organic compounds (VOCs) from six ground sites during September–December 2015. In

particular, this study combines CH₄ with four VOC species (ethane (C₂H₆), i-pentane (i-C₅H₁₂), n-pentane (n-C₅H₁₂), and toluene (C₇H₈)) for source partitioning of CH₄ emissions. Expanding on previous work by Jeong *et al.* [2016] and Ganesan *et al.* [2014], we developed a hierarchical Bayesian inversion (HBI) method for source partitioning. This novel approach incorporates both CH₄ and VOC measurements into a full Bayesian inverse analysis, allowing us to demonstrate the utility of multiple VOCs and a high-resolution spatially explicit emission model to separate natural gas and petroleum emissions from biological emissions.

2. Data and Methods

2.1. Atmospheric Measurements

We conducted an intensive measurement campaign including CH₄, selected VOCs, and carbon monoxide (CO) measurements from September to December 2015. Measurements of C₂H₆ have been used to constrain CH₄ emissions from the natural gas (NG) sector [e.g., Peischl *et al.*, 2013; McKain *et al.*, 2015], and we chose the other VOCs for the petroleum (PL) sector based on previous work [Kirchstetter *et al.*, 1996] where emission ratios of VOCs to CH₄ are available. The CH₄ and VOC measurements were made at five sites within SFBA and at a SFBA outflow site near Walnut Grove, CA (WGC, see Table S1 in the supporting information). With the exception of the Livermore site, CO flask sampling or continuous CO measurements were made at all sites. To capture CO measurements in the Livermore Valley, continuous measurements of CO were also collected at 27 m on a seventh tower (Sandia National Laboratories (SNL)) in the Livermore Valley for model-measurement comparison of CO mixing ratios. More details on the measurement method are provided in Text S1 in the supporting information [Colman *et al.*, 2001; Pétron *et al.*, 2012; BAAQMD, 2016].

2.2. Inverse Modeling

We expand on the HBI method developed by Jeong *et al.* [2016] combining both CH₄ and VOC measurements for source attribution. The details of the HBI method are explained in Jeong *et al.* [2016], and here we focus on the expanded equation that relates predictions to measurements to incorporate VOCs. We use the following linear model for estimating emissions [Zhao *et al.*, 2009; Jeong *et al.*, 2012a, 2012b, 2013, 2016; Wecht *et al.*, 2014]:

$$\mathbf{y} = \mathbf{K}\boldsymbol{\lambda} + \mathbf{v} \quad (1)$$

where \mathbf{y} is the background-subtracted measurement vector (see Text S1 for background [Jeong *et al.*, 2013]); $\mathbf{K} = \mathbf{F}\mathbf{E}$, \mathbf{F} is the footprint; \mathbf{E} is prior emissions; $\boldsymbol{\lambda}$ is a vector for scaling factors with a covariance matrix \mathbf{Q} ; and \mathbf{v} is a vector representing the model-measurement mismatch with a covariance matrix \mathbf{R} . The prior covariance matrix \mathbf{Q} represents our assumption of the confidence of prior emissions, and the diagonal components (i.e., σ_{λ}) of \mathbf{Q} are estimated during our inversion process instead of using fixed values (e.g., 50% uncertainty in Jeong *et al.* [2012a]). Following Jeong *et al.* [2016], we model σ_{λ} using a half Cauchy distribution with a scale parameter of 1. The parameters for the \mathbf{R} matrix are estimated by using the same exponential covariance function as in Jeong *et al.* [2016] (see Figure S6 in the supporting information for the estimates of the model-measurement uncertainty):

$$R_{i,j} = \eta^2 \exp\left(-\frac{1}{\tau} |t_i - t_j|\right) + \delta_{i,j} \sigma_{R_{s,x}}^2 \quad (2)$$

where η , τ , and $\sigma_{R_{s,x}}$ are parameters that define the covariance function, t is the measurement time, and δ is the Kronecker delta function. Expanding on Jeong *et al.* [2016], we add the components for VOCs to σ_R as shown in equation (2). The subscripts s and x in $\sigma_{R_{s,x}}$ indicate that σ_R is estimated as a mean value for each site (s) and gas species ($x = \{\text{CH}_4, \text{VOCs}\}$). This reflects the fact that model-measurement errors are not uniform across the sites and gas species. $\sigma_{R_{s,x}}$ is sampled from the half Cauchy distribution by using 50% of the mean measured signal as a scale parameter.

In this study we solve for $\boldsymbol{\lambda}$ at 0.1° resolution for SFBA after aggregating predictions, based on the high-resolution prior emissions (0.01° × 0.01°), and for regions outside SFBA we aggregate grid cells at the sub-region scale (i.e., California's air basins) in a similar way to that of Jeong *et al.* [2013, 2016]. We developed high-resolution (0.01°) spatially explicit prior emission (i.e., \mathbf{E}) maps for SFBA following the methods used in Jeong *et al.* [2012a, 2013, 2014, 2016], and the details for the prior emissions are described in Fischer and Jeong [2016]. The high-resolution prior emission maps for SFBA used in this study are shown in

Figure S1, and emission totals by sector (SFBA total=120 Gg/yr) are summarized in Table S3 where the sectors are categorized following recent regional inverse studies [e.g., Jeong *et al.*, 2013, 2016]. For the regions outside SFBA, we use the emission maps ($0.1^\circ \times 0.1^\circ$) from Jeong *et al.* [2016].

For footprint (**F**) calculations we use the coupled Weather Research and Forecasting and Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model [Skamarock *et al.*, 2008; Nehrkorn *et al.*, 2010] and adopt the setup used in Jeong *et al.* [2016]. In this setup, an ensemble of 500 STILT particles are run backward in time for 7 days driven with meteorology from the WRF model (version 3.6.1). Using WRF-STILT we simulate corresponding predictions to compare with the available measurements [Jeong *et al.*, 2012b]. The monthly mean footprint simulated by WRF-STILT is shown in Figure S2. To estimate the overall bias in the transport simulation, we compared measured and predicted CO (Figures S7 and S8) assuming that the potential bias in CO represents the bias in atmospheric transport for correcting the CH₄ emissions estimate [Jeong *et al.*, 2016]. However, we note that there is possibility that the overall uncertainty for the CH₄ emission estimate may be expanded or reduced because the errors in the CO emissions may be inversely correlated (i.e., anticorrelation) with errors in the transport model. The CO analysis suggests that predicted CO is consistent with the measurement within error yielding the best fit slope (predicted versus measured) of 0.93 ± 0.08 (at 68% confidence).

The linear model in equation (1) is expanded to solve for λ specific to biological (λ_{BIO}), natural gas (λ_{NG}), and petroleum (λ_{PL}) sources:

$$\begin{bmatrix} \mathbf{C}_{\text{CH}_4} & - & \mathbf{C}_{\text{CH}_4}^{\text{bg}} \\ \mathbf{C}_{\text{C}_2\text{H}_6} & - & \mathbf{C}_{\text{C}_2\text{H}_6}^{\text{bg}} \\ \mathbf{C}_{i\text{C}_5\text{H}_{12}} & - & \mathbf{C}_{i\text{C}_5\text{H}_{12}}^{\text{bg}} \\ \mathbf{C}_{nC_5\text{H}_{12}} & - & \mathbf{C}_{nC_5\text{H}_{12}}^{\text{bg}} \\ \mathbf{C}_{\text{C}_7\text{H}_8} & - & \mathbf{C}_{\text{C}_7\text{H}_8}^{\text{bg}} \end{bmatrix} = \begin{bmatrix} \mathbf{FE}_{\text{BIO}} & \mathbf{FE}_{\text{NG}} & \mathbf{FE}_{\text{PL}} \\ \mathbf{0} & \mathbf{FE}_{\text{NG}} f_{\text{C}_2\text{H}_6\text{NG}}^* & \mathbf{FE}_{\text{PL}} f_{\text{C}_2\text{H}_6\text{PL}}^* \\ \mathbf{0} & \mathbf{0} & \mathbf{FE}_{\text{PL}} f_{i\text{C}_5\text{H}_{12}\text{PL}}^* \\ \mathbf{0} & \mathbf{0} & \mathbf{FE}_{\text{PL}} f_{nC_5\text{H}_{12}\text{PL}}^* \\ \mathbf{0} & \mathbf{0} & \mathbf{FE}_{\text{PL}} f_{\text{C}_7\text{H}_8\text{PL}}^* \end{bmatrix} \begin{bmatrix} \lambda_{\text{BIO}} \\ \lambda_{\text{NG}} \\ \lambda_{\text{PL}} \end{bmatrix} + \mathbf{v} \quad (3)$$

where \mathbf{C}_X is the measurement vector ($n \times 1$) for gas species X , \mathbf{C}_X^{bg} is the background vector ($n \times 1$), and f_X^* is the emission ratio of VOC X to CH₄. For VOC background, we fit a simple linear model to measurements at Trinidad Head (THD, gas analysis by NOAA) to generate continuous VOC background mixing ratios (Figure S3). For example, given a few available data points the measured C₂H₆ from THD shows a linear trend during the study period, and a similar linear trend is shown at a global background site in Iceland [Helmig *et al.*, 2016]. Because this linear model could introduce mean bias, equation (3) is modified to incorporate the mean adjustment as

$$\mathbf{y} = \mathbf{K}\lambda + \mathbf{D} + \mathbf{v} \quad (4)$$

where $\mathbf{K}\lambda$ represents the first two terms on the right-hand side of equation (3) and \mathbf{D} is a vector for mean adjustments, which is estimated during the hierarchical inverse process. Each element of \mathbf{D} represents the mean background adjustment and other potential biases (e.g., measurement offsets and transport biases) for each combination of sites and species. The estimated mean biases are provided in Table S2 and Figure S4. Although \mathbf{D} is estimated with a large uncertainty bound for each combination of sites and gas species (see Table S2 and Figure S4), there exists some variability in estimated \mathbf{D} across the sites and species. This suggests that for a subset of the sites and species there are undiagnosed residual mean biases (e.g., due to uncaptured subgrid-scale transport error) that are often found in the comparison between posterior predictions and measurements [e.g., Johnson *et al.*, 2016]. The small estimated bias at the Sutro (STR) site suggests that on average the measurements at THD are representative of background concentrations for the SFBA region. For the matrix that constitutes predictions (i.e., \mathbf{K}), \mathbf{FE}_{BIO} ($n \times k$ matrix; k is the number of pixels (0.1°) or sub-regions) is the predicted mixing ratio for biological (BIO) sources (i.e., all sectors including landfill, livestock, and wastewater except for the NG and PL sectors), \mathbf{FE}_{NG} ($n \times k$) is the predicted mixing ratio for NG, and \mathbf{FE}_{PL} ($n \times k$) is the predicted mixing ratio for PL. Zero is a zero matrix with a dimension of $n \times k$. Applying this inversion scheme, we perform inversions for two separate seasons of early fall (September–October) and late fall (November–December), and estimated emissions are summarized for the combined season.

Following the HBI method in Jeong *et al.* [2016], we solve equation (4) by using a Markov chain Monte Carlo (MCMC) technique (see Jeong *et al.* [2016] for details), and here we add additional parameters for \mathbf{D} (equation (4)) and f_{χ}^* . Expanding on Jeong *et al.* [2016], we estimate the following joint parameter set:

$$\Theta = \{\lambda, \mu_{\lambda}, \sigma_{\lambda}, \sigma_{\mathbf{R}}, \eta, \tau, f_{\chi}^*, \mathbf{D}\}$$

where μ_{λ} is the prior (i.e., hyper-parameter) mean for λ and σ_{λ} is the uncertainty for λ . When all four VOCs are used, the total number of parameters to be estimated is 1822 (585 for each of $\{\lambda, \mu_{\lambda}, \sigma_{\lambda}\}$, 30 for each of $\{\sigma_{\mathbf{R}}, \mathbf{D}\}$, 5 for f_{χ}^* , and a single estimate for each of $\{\tau, \eta\}$). We note that the hierarchical approach is a sensible choice when there are more parameters than data because the additional structure through the hierarchy reduces the overfitting problem [Gelman *et al.*, 2014].

To build MCMC samplers (sample size = 20,000), the NIMBLE package (version 0.5; [de Valpine *et al.*, 2016]) is used together with the R statistical language (<https://cran.r-project.org/>). NIMBLE extends the Bayesian inference Using Gibbs Sampling [Gilks *et al.*, 1994] language and is flexible to implement complex model specifications as in our study. We use a Metropolis-Hastings adaptive random-walk sampler available in NIMBLE, and the proposal distribution is automatically adapted to a target acceptance rate (e.g., 0.4) depending on the sampler [de Valpine *et al.*, 2016] to ensure optimal MCMC mixing [Roberts *et al.*, 1997]. For VOC:CH₄ ratios (i.e., f_{χ}^*), for example, $f_{\text{C}_2\text{H}_6\text{NG}}^*$ and $f_{\text{C}_2\text{H}_6\text{PL}}^*$ are the C₂H₆:CH₄ emission ratios for the NG and PL sectors, respectively, which are optimized based on the VOC measurements during the inversion. For $f_{\text{C}_2\text{H}_6\text{NG}}^*$, the median value from the posterior samples (Figure S5) is similar to the prior assumption of 0.04, which is a typical value for C₂H₆:CH₄ reported for the SFBA NG distribution system [PG&E, 2015]. The median value for the posterior estimate for $f_{\text{C}_2\text{H}_6\text{PL}}^*$ is also similar to the prior assumption of 0.07 (see Table S4 for the prior values). For other VOCs, the estimated VOC:CH₄ ratios from HBI (Figure S5) are consistent with the prior values except for C₇H₈, which shows somewhat lower posterior estimates (median = 0.12) for its VOC:CH₄ ratio compared to the prior value (0.16).

3. Results

3.1. CH₄ Emissions in SFBA

Our HBI analysis estimates that SFBA's CH₄ emissions (including wetland, 3% of the anthropogenic total in our prior emission) are 174–282 Gg CH₄/yr (median = 225, at 95% confidence), higher (1.4–2.2 times) than a recent inventory (126 Gg CH₄/yr in 2011) from BAAQMD [BAAQMD, 2015]. A simple regression analysis of model predictions versus measurements also indicates that actual CH₄ emissions are more than 2 times higher than those from the prior emission model (Figure S9).

In comparison with previous work on the SFBA CH₄ emissions, we note that the estimated total CH₄ emissions are consistent with results from independent studies by Fairley and Fischer [2015] and Jeong *et al.* [2016] which reported 180–300 Gg CH₄/yr (at 95% confidence) and 159–340 Gg (at 95% confidence), respectively. Here we note that Fairley and Fischer [2015] estimated CH₄ emissions in a decade-long CH₄:CO correlation analysis from surface air-quality monitoring stations, while Jeong *et al.* [2016] estimated the total CH₄ emissions in a yearlong atmospheric inversion study without source partitioning, driven primarily by CH₄ measurements from three SFBA tower sites (STR, SNL, and WGC). Jeong *et al.* [2016] used different prior emissions (0.1° resolution) with 20% higher total for prior SFBA emissions and conducted inversions at a coarser resolution of 0.3°. In addition, Wecht *et al.* [2014] obtained a range of central estimates for SFBA CH₄ emissions (370, 390, and 430 Gg/yr) from an inverse model analysis of airborne CH₄ measurements obtained during the May–June 2010 California Nexus campaign. While Wecht *et al.* [2014] do not report uncertainty bounds on these estimates, the minimum of their range falls outside the 95% confidence interval of the current estimate and is also roughly double the central estimates of Fairley and Fischer [2015] and Jeong *et al.* [2016].

Figure 1a shows the median posterior emissions after inversion by sector, adjusting the prior emissions for individual 0.1° pixels based on the corresponding posterior scaling factors [Jeong *et al.*, 2013, 2016; Wecht *et al.*, 2014]. Among individual sectors, in particular, our HBI analysis suggests that the posterior landfill (LF) emissions are much higher than those of the prior, although the ratio of LF to total (51%) in the posterior estimate is similar to the prior one (54%). We note that source partitioning using this method depends on the accuracy of the relative source composition for each pixel in the prior emission

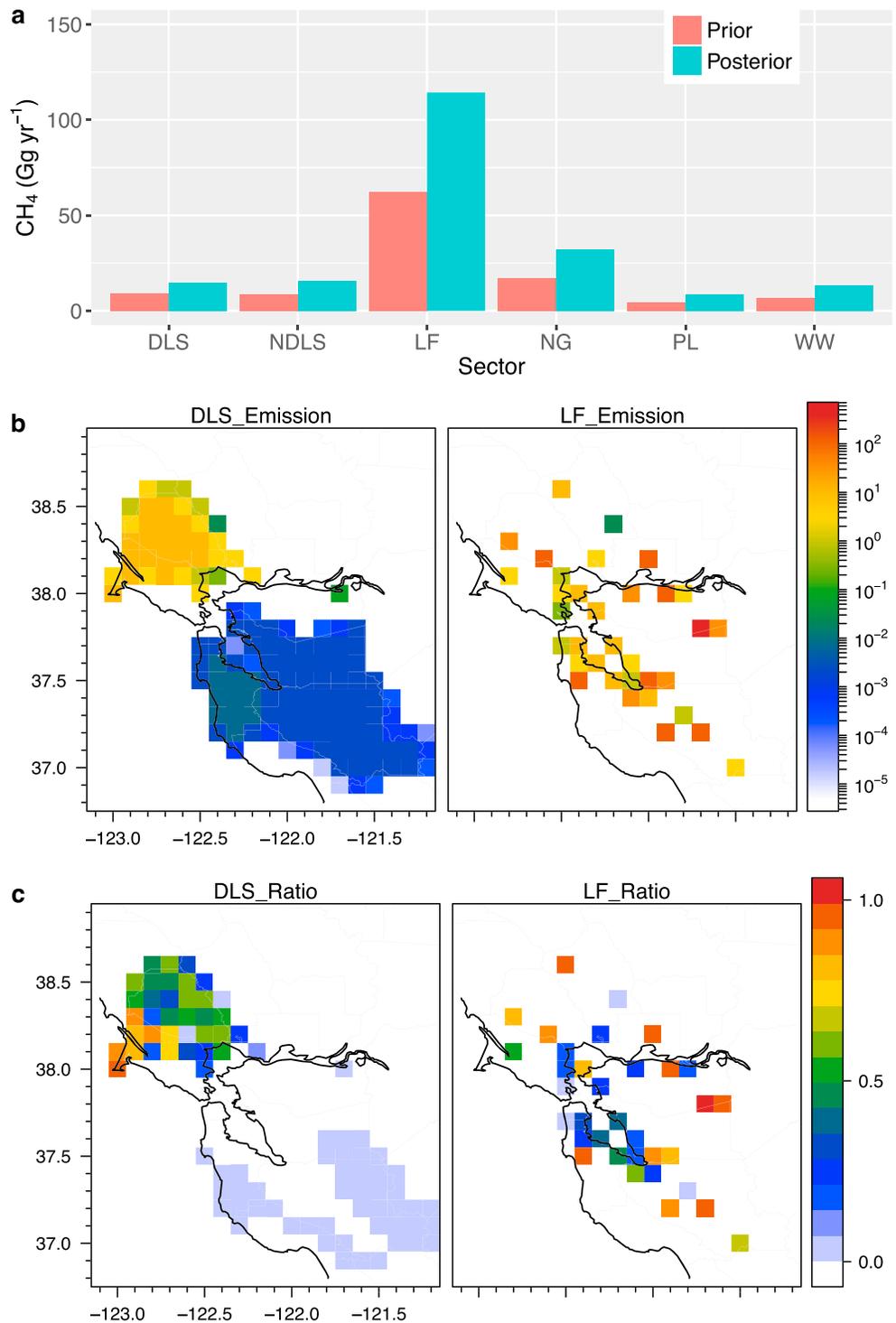


Figure 1. (a) CH₄ emissions (only posterior median emissions shown) in SFBA by sector using CH₄ measurements only, (b) dairy livestock (DLS) and landfill (LF) emission flux (nmol/m²/s) maps at 0.1° × 0.1° resolution, (c) and maps of ratios of DLS and LF to the total emissions for each 0.1° pixel. The sectors in the bar plot in Figure 1a includes dairy livestock (DLS), nondairy livestock (NDLS), landfill (LF), natural gas (NG), petroleum refining and on-road mobile sources (PL), and wastewater (WW). For this result, we assume that the ratios of individual sector emissions to the total for each 0.1° pixel is known from the prior CH₄ emission map. Then we scaled each sector prior within each pixel by the corresponding posterior scaling factor to obtain posterior emissions. Note that the inversion is performed at 0.1° (~10 km) resolution for SFBA aggregating predicted mixing ratio concentrations based on the prior emissions at 0.01° (~1 km) resolution. In Figures 1b and 1c, the longitude labels (x axis) are the same for all four figures.

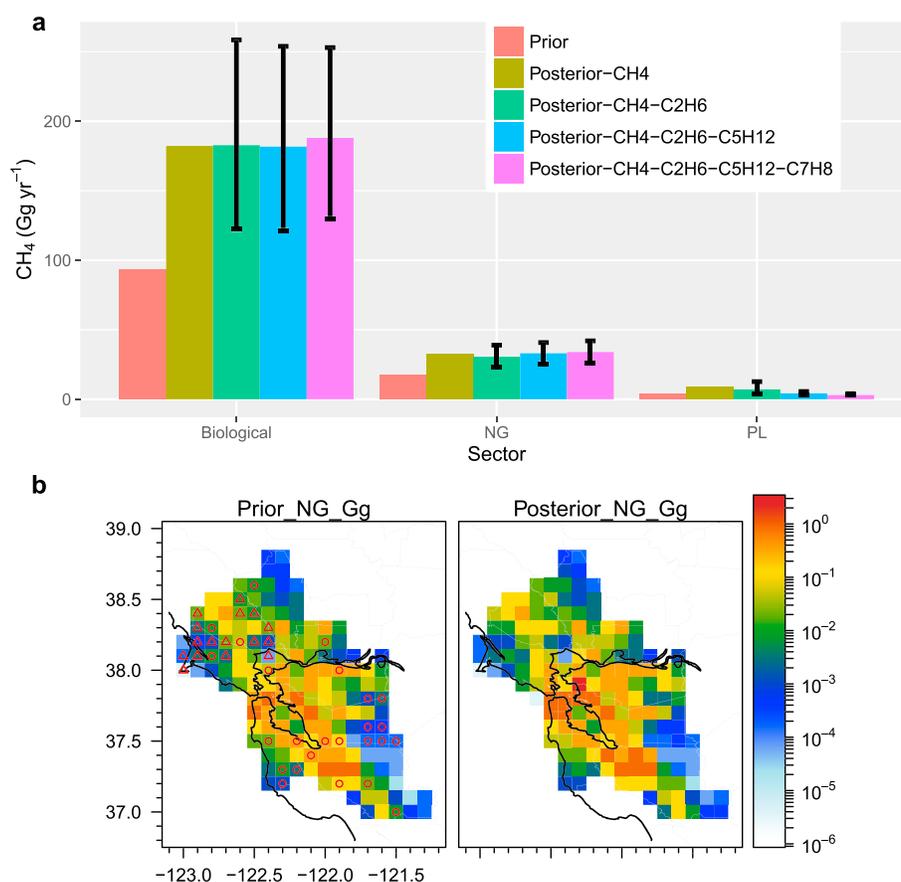


Figure 2. (a) Comparison of posterior CH₄ emissions (at 95% confidence) by source sector using CH₄, C₂H₆, i-C₅H₁₂, n-C₅H₁₂, and C₇H₈ measurements and (b) comparison of prior and posterior natural gas emissions (Gg CH₄/yr; 0.1° resolution) for SFBA. In Figure 2a, the estimate using CH₄ only is shown without confidence intervals as in Figure 1. In the prior emission map in Figure 2b, the red open circles and triangles indicate the pixels where landfill and dairy emissions are dominant, respectively (>50% of the pixel total). The natural gas emissions are small in most of the pixels where landfill and dairy emissions are dominant.

model and may not be completely independent; for this reason, we do not show confidence intervals on the values shown in this figure.

A further analysis of the spatial distribution of emissions suggests that inversions using a high resolution as in this study can be useful to separate emissions for some source sectors from those of other sources. Figures 1b and 1c show the LF emission map and the ratio of LF to total emissions for each 0.1° pixel in SFBA. Many of the pixels, which have high emission fluxes, also show high ratios of LF to the total. Summing emissions from the pixels with greater than 0.8 in the LF ratio relative to the total, we find that the total from those high-emitting pixels accounts for 85% of the total LF emissions in SFBA (Figure S10). We can apply a similar analysis to the dairy livestock (DLS) sector, which has been reported to emit more CH₄ emissions than suggested by the state inventory in California [Jeong *et al.*, 2013, 2016; Johnson *et al.*, 2014, 2016] although in SFBA the difference between the prior and posterior DLS emissions is less noticeable than the LF sector. As shown in Figures 1b and 1c, the dairy emission sources in SFBA are concentrated in Marin and Sonoma counties, and in the majority of those pixels with DLS emission sources the ratios of DLS to the total are greater than 0.5.

3.2. Source Apportionment Using VOCs

Combining CH₄ and VOC measurements, we use a more robust method for source attribution than we applied in the previous section and separate CH₄ emissions into BIO, NG, and PL sectors. Figure 2a shows the estimated emissions by sector considering different sets of VOCs. We first performed the inversion with CH₄ and C₂H₆ only, and then added i-C₅H₁₂ and n-C₅H₁₂ to further constrain the PL sector. A third inversion

was conducted combining CH₄ and all four VOCs. When i-C₅H₁₂, n-C₅H₁₂, and C₇H₈ are used in addition to C₂H₆, the posterior uncertainty for PL emissions are significantly reduced compared to the case with C₂H₆ only. As expected, the BIO and NG emissions do not vary much across different inversions because the same CH₄ and C₂H₆ measurements are used for all inversions. The uncertainty for the BIO emission is slightly reduced when the PL sector is further constrained by addition of more VOCs. The inversion results using CH₄ and all four VOC measurements suggest that SFBA CH₄ emissions (including wetlands, ~3% of total) are 166–289 Gg CH₄/yr (median = 226, at 95% confidence), which are 1.3–2.3 times higher than a recent inventory (126 Gg CH₄/yr) from BAAQMD. This estimate is very similar to the one estimated by using CH₄ measurements only (174–282 Gg). We note that if we incorporate the additional uncertainty of 8% (about the median estimate) based on the CO measurement-prediction comparison assuming the CO analysis is applicable to CH₄ [Jeong *et al.*, 2016] (section 2.2 and Figure S8), the estimated CH₄ emissions are 154–297 Gg CH₄/yr (1.2–2.4 times the BAAQMD inventory, 95% confidence). We also note that undiagnosed sources of uncertainty (e.g., unresolved transport errors) may increase these error bounds beyond that indicated here. For example, there are cases where the simulated footprints over SFBA are too narrow, not spreading over the region, and local emissions may not be captured properly.

Attributing the total CH₄ emissions to different sources, we estimate that the BIO, NG, and PL emissions in SFBA account for 82, 15, and 2% of the total (remaining 1% from other miscellaneous sources), respectively, suggesting that the BIO emissions are much higher than fossil-fuel-based emissions. Similarly, our prior emission model suggests that the BIO source was dominant (78%) in SFBA. The HBI analysis estimates that the posterior emission for the BIO sector is 126–249 Gg CH₄/yr (at 95% confidence), higher than the prior (94 Gg CH₄/yr) by factors of 1.3–2.6.

We estimate that total NG CH₄ emissions for SFBA using all VOCs are 26–42 Gg CH₄/yr, 1.5–2.4 times higher than the prior (17 Gg CH₄/yr; Figure 2a). In Figure 2a, the NG emissions estimated from the CH₄-only inversion are consistent with those estimated by using observations of both CH₄ and VOCs. This is likely due to the fact that most of the NG-dominant pixels are spatially separated from the major BIO sources (i.e., landfill and dairy emissions; Figures 1c and 2b) and the high-resolution inversion in this study distinguishes the majority of NG emissions from the landfill and dairy emissions. Assessment of multicollinearity (a measure for correlation between two or more variables) between BIO, NG, and PL sources also suggests that the three sources are not strongly correlated. To check multicollinearity between BIO, NG, and PL sources, using BIO, NG, and PL predictions, we calculated the variance inflation factors (VIF) which were 2–4. In general, the VIF value larger than 10 indicates that collinearity can be problematic in a linear model [Myers, 1990]. However, we note that the inversion may not completely separate emissions between individual pixels because of anticorrelation [Jeong *et al.*, 2012a, 2013]. The spatial distribution of CH₄ emissions for the NG sector is compared between prior and posterior estimates in Figure 2b, where emissions in some pixels are elevated after inversion. According to our prior model, the vast majority (99%) of the NG emissions in SFBA is from the NG distribution sector (Table S3), and the elevated emissions in the posterior NG emission map indicate that the prior emissions for the NG distribution sector are underestimated.

Our posterior estimates for NG CH₄ emissions are comparable to that estimated from a recent study aimed at estimating spatially explicit natural gas emissions across California based on a bottom-up approach [Jeong *et al.*, 2014]. Jeong *et al.* [2014] estimates a total of 30 Gg CH₄/yr for SFBA from the natural gas sector, 93% of which are from the NG transmission and distribution subsectors. In comparison, the total for NG in our prior emission model is 17 Gg CH₄/yr, which is based on BAAQMD's inventory. Our posterior estimate for NG is 0.87–1.4 times the total in Jeong *et al.* [2014]. Overall, our posterior estimate for the NG sector supports the finding of Jeong *et al.* [2014] that their SoCAB total emission arising from NG distribution was lower than the top-down estimate (for the year 2010) [Peischl *et al.*, 2013] by a factor of ~2. In the inversion incorporating data for all four VOCs, we estimate that the posterior PL emission is 3–4 Gg CH₄/yr (at 95%), which is consistent with the prior (4 Gg CH₄).

4. Discussion and Conclusions

The multitracer method used in this study effectively separates BIO source CH₄ emissions from NG and PL source emissions in SFBA. This study estimates that 82% (i.e., median ratio of 185/226) of the total emission

in SFBA is from biological sources, suggesting that the ratio of biological versus fossil emissions in SFBA is different from that of SoCAB, where multiple studies suggest that the proportion of fossil-fuel CH_4 is considerably higher (58–65% [Hopkins *et al.*, 2016], ~70% [Wennberg *et al.*, 2012], and 41–68% [Peischl *et al.*, 2013]). Similarly, in the California Greenhouse Gas Emission Measurements (CALGEM) prior model [Jeong *et al.*, 2016], which is scaled to match the 2012 CARB state inventory [CARB, 2014], the fossil-fuel sources for SoCAB account for 35% of the total (349 Gg CH_4/yr), which is higher than that (22%) of our prior model for SFBA (Table S3). Among the BIO sources, our analysis shows that the landfill sector is underestimated in the prior emission, contributing significantly to the large discrepancy between the prior and posterior emissions estimates. Our result for SFBA is different from that for SoCAB where Peischl *et al.* [2013] suggested that landfill emissions in SoCAB are generally consistent with CARB's inventory estimate.

Furthermore, the VOC-based analysis also indicates that emissions from the fossil-fuel sources in SFBA are relatively small compared to the biological sources. These results suggest that landfill emissions are likely a dominant source of the total emissions in SFBA. To further investigate the discrepancy in the landfill emissions in SFBA, we analyzed landfill data sets from U.S. Environmental Protection Agency (EPA)'s official inventory [U.S. Environmental Protection Agency, 2016] and U.S. EPA's Greenhouse Gas Reporting Program (GHGRP, <https://ghgdata.epa.gov/ghgp/main.do>, accessed June 2016). We find that landfill emissions from EPA are higher than our prior estimate based on CARB's inventory or BAAQMD's estimate, supporting the result from our inverse analysis (Figure S11). An independent estimate [Maasakkers *et al.*, 2016] based on EPA's official national inventory also suggests that California's landfill total emissions from EPA's inventory (for 2012) are higher (~500 Gg) than that (~340 Gg) of CARB [CARB, 2014]. When we revise our prior inventory using EPA's landfill emissions and the NG emissions from Jeong *et al.* [2014], the revised inventory is marginally consistent with our posterior estimates for SFBA's total CH_4 emissions (Figure S11). Note that for the revision of the landfill prior we scaled the GHGRP landfill emissions to match EPA's official inventory considering that GHGRP includes only ~70% of the total landfill emissions reported in EPA's inventory. The scaled landfill for California (in 2012) is ~500 Gg, similar to an independent estimate using the EPA inventory [Maasakkers *et al.*, 2016]. More studies using facility-level measurements of CH_4 and relevant VOCs (e.g., landfill flux measurements) would be useful to verify the results of this study. Recent studies reported that livestock is the main source of CH_4 emissions in California and likely is underestimated in the CARB inventory [Jeong *et al.*, 2012a, 2013, 2016; Johnson *et al.*, 2014, 2016; Wecht *et al.*, 2014]. Because we derived the emission factors for dairy livestock based on CARB's inventory, it is possible that the dairy emissions in SFBA are also higher than the prior estimate. This explanation for the underestimation of the biological sources in the prior emission model merits further verification.

The inversion result based on VOC measurements suggests that the NG CH_4 emission rate in SFBA is relatively lower than other urban regions in the United States. Natural gas consumption in SFBA shows seasonality with winter having the highest consumption (~2 times that of summer). If we assume that CH_4 emissions for the NG distribution sector are proportional to NG consumption and apply the seasonal consumption relative to that for our study period to our posterior emissions (26–42 Gg CH_4), we estimate that annual NG emissions for SFBA are 23–38 Gg CH_4 (95% confidence; mostly related to NG distribution). These annual estimates are equivalent to 0.3–0.5% (in mass ratio) of the total NG consumption, which is estimated to be $1.26 \times 10^{10} \text{ m}^3$ for 2011 in BAAQMD's official inventory [BAAQMD, 2015]. In SoCAB, the NG emissions are estimated to be 0.7–3% of the total NG consumption; in SoCAB, however, separating distribution NG emissions from other NG sources is more challenging than in SFBA because of oil and gas production and geologic seeps [Wennberg *et al.*, 2012; Peischl *et al.*, 2013]. Farrell *et al.* [2013] estimate up to ~60 Gg CH_4/yr from the La Brea Tar Pits alone. If more CH_4 emissions are attributable to local geologic seeps, these percentage emission rates relative to NG consumption would decrease. McKain *et al.* [2015] reported that the emission rate in the Boston urban area is $2.7 \pm 0.6\%$ (at 95% confidence) of NG consumption.

In summary, we presented a Bayesian inverse analysis based on a combination of a high-resolution prior emission model and both CH_4 and VOC measurements separating fossil-fuel source emissions from those of biological sources. The estimated total emissions from this study are consistent with recent independent studies [Fairley and Fischer, 2015; Jeong *et al.*, 2016], suggesting that the inverse framework based on the measurement network can be an effective approach to monitoring long-term spatial and temporal changes in SFBA emissions. However, it is possible that undiagnosed sources of error (e.g., unresolved transport errors)

affect the CH₄ emissions estimates. In the future, a combination of improved prior emissions and meteorological models, expanded multigas measurements (e.g., additional VOC measurements), and inverse model analyses should reduce uncertainty in SFBA's GHG emissions. In SFBA and across the state, where livestock emissions are important, additional measurements of VOCs (e.g., methanol as in *Guha et al.* [2015]) could be incorporated into the inversion system developed in this study to separate dairy emissions from other emissions sources. Increased practice of combining facility-specific emission measurements and regionally representative measurements of source-specific tracers (e.g., CO, VOCs, and potentially CH₄ isotopes) [Townsend-Small et al., 2012; Peischl et al., 2013; Guha et al., 2015] would be very useful given the importance of distinguishing dominant CH₄ sources for prioritizing mitigation of large-scale events, such as the well failure at Aliso Canyon [Conley et al., 2016].

Acknowledgments

Authors acknowledge BAAQMD staff in the Planning and Climate Protection Division for assistance with emissions inventory development and staff in the Meteorology, Measurements and Rules Division for assistance with air quality data collection and site access. The methane and ethane data used in the inversion, THD background data, and high-resolution prior emissions are in supplements, and the CALGEM prior emission distribution is available at <http://calgem.lbl.gov/>. This analysis was supported by the California Energy Commissions Public Interest Environmental Research program, with work at LBNL conducted under U.S. Department of Energy contract DE-AC02-05CH11231.

References

- BAAQMD (2015), Bay Area Emissions Inventory Summary Report: Greenhouse Gases Base Year 2011. [Available at http://www.baaqmd.gov/~media/files/planning-and-research/emission-inventory/by2011_ghgsummary.pdf, updated January 2015]; accessed March 2016.]
- BAAQMD (2016), 2015 Air Monitoring Network Plan. [Available at http://www.baaqmd.gov/~media/files/technical-services/2015_network_plan.pdf?la=en; accessed May 2016.]
- California Legislative Information (2014), Senate Bill No. 605 Short Lived Climate Pollutants, California Senate Bill. [Available at http://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=2013201405SB605(accessed July 2016).]
- CARB (2014), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report, Accessed January 2015. [Available at <http://www.arb.ca.gov/cc/inventory/inventory.htm>, version March 2014].]
- CARB (2015), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report, Accessed September 2015. [Available at <http://www.arb.ca.gov/cc/inventory/inventory.htm>, version April 2015].]
- Colman, J. J., A. L. Swanson, S. Meinardi, B. C. Sive, D. R. Blake, and F. S. Rowland (2001), Description of the analysis of a wide range of volatile organic compounds in whole air samples collected during PEM-tropics A and B, *Anal. Chem.*, **73**, 3723–3731, doi:10.1021/ac10027g.
- Conley, S., G. Franco, I. Faloona, D. R. Blake, J. Peischl, and T. B. Ryerson (2016), Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA, *Science*, doi:10.1126/science.aaf2348.
- Cui, Y. Y., et al. (2015), Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin, *J. Geophys. Res. Atmos.*, **120**, 6698–6711, doi:10.1002/2014JD023002.
- de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik (2016), Programming with models: Writing statistical algorithms for general model structures with NIMBLE, *J. Comput. Graphical Stat.*, doi:10.1080/10618600.2016.1172487.
- Fairley, D., and M. L. Fischer (2015), Top-down methane emissions estimates for the San Francisco Bay Area from 1990 to 2012, *Atmos. Environ.*, doi:10.1016/j.atmosenv.2015.01.065.
- Farrell, P., D. Culling, and I. Leifer (2013), Transcontinental methane measurements: Part 1. A mobile surface platform for source investigations, *Atmos. Environ.*, doi:10.1016/j.atmosenv.2013.02.014.
- Fischer, M. L., and S. Jeong (2016), Evaluating the Bay Area methane emission inventory, BAAQMD report (2014-108). [Available at http://www.baaqmd.gov/~media/files/planning-and-research/emission-inventory/baaqmd-2014-108-sfba-ch4-emissions_20160330.pdf?la=en]
- Ganesan, A. L., et al. (2014), Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, *Atmos. Chem. Phys.*, **14**, 3855–3864, doi:10.5194/acp-14-3855-2014.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2014), Bayesian Data Analysis, 3rd ed., Chapman & Hall/CRC, Boca Raton, Fla.
- Gilks, W. R., A. Thomas, and D. J. Spiegelhalter (1994), A language and program for complex Bayesian modeling, *J. R. Stat. Soc. Ser. D*, **43**(1), 169–177.
- Guha, A., D. R. Gentner, R. J. Weber, R. Provencal, A. Gardner, and A. H. Goldstein (2015), Source apportionment of methane and nitrous oxide in California's San Joaquin Valley at CalNex 2010 via positive matrix factorization, *Atmos. Chem. Phys.*, **15**, 12,043–12,063, doi:10.5194/acp-15-12043-2015.
- Helmig, D., et al. (2016), Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production, *Nat. Geosci.*, **9**, doi:10.1038/NGEO2721.
- Hopkins, F. M., E. A. Kort, S. E. Bush, J. R. Ehleringer, C.-T. Lai, D. R. Blake, and J. T. Randerson (2016), Spatial patterns and source attribution of urban methane in the Los Angeles Basin, *J. Geophys. Res. Atmos.*, **121**, 2490–2507, doi:10.1002/2015JD024429.
- Hsu, Y.-K., T. VanCuren, S. Park, C. Jakober, J. Herner, M. FitzGibbon, D. R. Blake, and D. D. Parrish (2010), Methane emissions inventory verification in southern California, *Atmos. Environ.*, **44**, 1–7, doi:10.1016/j.atmosenv.2009.10.002.
- Hu, A., Y. Xu, C. Tebaldi, W. M. Washington, and V. Ramanathan (2013), Mitigation of short-lived climate pollutants slows sea-level rise, *Nat. Clim. Change*, **3**, 730–734, doi:10.1038/nclimate1869.
- Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012a), Seasonal variation of CH₄ emissions from central California, *J. Geophys. Res.*, **117**, D11306, doi:10.1029/2011JD016896.
- Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012b), Seasonal variations in N₂O emissions from central California, *Geophys. Res. Lett.*, **39**, L16805, doi:10.1029/2012GL052307.
- Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak, and M. L. Fischer (2013), A multitower measurement network estimate of California's methane emissions, *J. Geophys. Res. Atmos.*, **118**, 11,339–11,351, doi:10.1002/jgrd.50854.
- Jeong, S., D. Millstein, and M. L. Fischer (2014), Spatially explicit methane emissions from petroleum production and the natural gas system in California, *Environ. Sci. Technol.*, **48**, 5982–5990.
- Jeong, S., et al. (2016), Estimating methane emissions in California's urban and rural regions using multi-tower observations, *J. Geophys. Res. Atmos.*, **121**, 13,031–13,049, doi:10.1002/2016JD025404.
- Johnson, M. S., E. L. Yates, L. Iraci, M. Loewenstein, J. Tadić, K. J. Wecht, S. Jeong, and M. L. Fischer (2014), Analyzing source apportioned methane in northern California during Discover-AQ-CA using airborne measurements and model simulations, *Atmos. Environ.*, **99**, 248–256, doi:10.1016/j.atmosenv.2014.09.068.

- Johnson, M. S., X. Xi, S. Jeong, E. L. Yates, L. T. Iraci, T. Tanaka, M. Loewenstein, J. M. Tadić, and M. L. Fischer (2016), Investigating seasonal methane emissions in Northern California using airborne measurements and inverse modeling, *J. Geophys. Res. Atmos.*, *121*, 13,753–13,767, doi:10.1002/2016JD025157.
- Kirchstetter, T. W., B. C. Singer, R. A. Harley, G. R. Kendall, and W. Chan (1996), Impact of oxygenated gasoline use on California light-duty vehicle emissions, *Environ. Sci. Technol.*, *30*, 661–670.
- Maasakkers, J. D., et al. (2016), Gridded national inventory of U.S. methane emissions, *Environ. Sci. Technol.*, *50*, 13,123–13,133, doi:10.1021/acs.est.6b02878.
- McKain, K., et al. (2015), Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts, *Proc. Natl. Acad. Sci. U.S.A.*, *112*(7), doi:10.1073/pnas.1416261112.
- Myers, R. H. (1990), *Classical and modern regression with applications*, 2nd ed., Duxbury, Boston, Mass.
- Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled Weather Research and Forecasting–Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model, *Meteorol. Atmos. Phys.*, *107*(1), 51–64, doi:10.1007/s00703-010-0068-x.
- Peischl, J., et al. (2013), Quantifying sources of methane using light alkanes in the Los Angeles basin, California, *J. Geophys. Res. Atmos.*, *118*, 4974–4990, doi:10.1002/jgrd.50413.
- Pétron, G., et al. (2012), Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study, *J. Geophys. Res.*, *117*, D04304, doi:10.1029/2011JD016360.
- PG&E (2015), Online reporting of natural gas composition. [Available at http://www.pge.com/pipeline/operations/gas_quality/index.page; accessed Jan-Dec, 2015.]
- Roberts, G. O., A. Gelman, and W. R. Gilks (1997), Weak convergence and optimal scaling of random walk Metropolis algorithms, *Ann. Appl. Probab.*, *7*, 110–120, doi:10.1214/aoap/1034625254.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, X. Z. Huang, W. Wang, and J. G. Powers (2008), A description of the advanced research WRF version 3 Technical Note 475 + STR. Mesoscale and Microscale Meteorology Division, NCAR, Boulder, Colo.
- Townsend-Small, A., S. C. Tyler, D. E. Pataki, X. Xu, and L. E. Christensen (2012), Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of “fugitive” fossil fuel emissions, *J. Geophys. Res.*, *117*, D07308, doi:10.1029/2011JD016826.
- Turner, A. J., et al. (2015), Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmos. Chem. Phys.*, *15*, 7049–7069, doi:10.5194/acp-15-7049-2015.
- US Environmental Protection Agency (2016), Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2014. EPA 430-R-16-002. [Available at <http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html>, accessed May 2016.]
- Wecht, K. J., D. J. Jacob, M. P. Sulprizio, G. W. Santoni, S. C. Wofsy, R. Parker, H. Bösch, and J. Worden (2014), Spatially resolving methane emissions in California: Constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI, geostationary) satellite observations, *Atmos. Chem. Phys.*, *14*, 8173–8184, doi:10.5194/acp-14-8173-2014.
- Wennberg, P. O., et al. (2012), On the sources of methane to the Los Angeles atmosphere, *Environ. Sci. Technol.*, *46*(17), 9282–9289, doi:10.1021/es301138y.
- Wong, K. W., D. Fu, T. J. Pongetti, S. Newman, E. A. Kort, R. Duren, Y.-K. Hsu, C. E. Miller, Y. L. Yung, and S. P. Sander (2015), Mapping CH₄:CO₂ ratios in Los Angeles with CLARS-FTS from Mount Wilson, California, *Atmos. Chem. Phys.*, *15*, 241–252, doi:10.5194/acp-15-241-2015.
- Wunch, D., et al. (2016), Quantifying the loss of processed natural gas within California’s South Coast Air Basin using long-term measurements of ethane and methane, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-359.
- Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, *J. Geophys. Res.*, *114*, D16302, doi:10.1029/2008JD011671.